Food Markets' Structural Empirical Analysis: A Review of Methods and Topics

Celine Bonnet, Carly Trachtman, Molly Van Dop, and Sofia B. Villas-Boas *

October 30, 2018

Abstract

We review the current literature on applied policy analysis and empirical structural food industrial organization research. Within that context, we provide an overview on the state of research and summarize the core problems researchers face when estimating the demand and structural supply models implemented in applied research to date. We focus on important themes, such as providing a better understanding of vertical relationships in food markets, price formation, competition issues, environmental, and nutritional policies.

Keywords: Structural demand models, Structural supply models of vertical supply chains, Structural empirical methods, Food markets, Policy analysis.

JEL Classification: M30, Q18, Q25, Q21, C12, C24.

1 Introduction

The analysis of food markets is of utmost importance, as food expenditures represent 13% of the total budget of households in the European Union (EU) in 2014, according to the Statistical Office of the EU. Moreover, this is the largest manufacturing sector in the EU, as the food and drink industry reached a turnover of 3.9 trillion Euros in 2014 and represents 10% of all EU employment. In addition, food markets are critical to competition policy analysis. Concentration in the food processing industry and retail sectors is much higher than in agricultural market. Whereas the C5 concentration ratio in agriculture in 2010 accounted

*We thank the editor Salvatore Di Falco and an anonymous reviewer for their invaluable suggestions as well as Peter Berck for helpful comments. Bonnet: Toulouse School of Economics, INRA, University of Toulouse I Capitole; Trachtman, Van Dop, Villas-Boas: Department of Agricultural and Resource Economics, University of California, Berkeley, CA, 94720. Corresponding author: celine.bonnet@tse-fr.eu.

for 0.19%, the market share of the top five firms in the EU food industry was at an average of 56% in 2012 in 14 of the EUs member states and the share of the top five retailers exceeded 60% in 13 member states at the same time. Concentration helps achieve economies of scale but also endows higher bargaining power that could lead to unfair practices. The emergence of store brands gives rise to a series of issues such as innovation incentives, consumer choices, and price levels. The regulatory environment, such as product market regulation, administrative regulation, domestic economic regulation, and public health policies largely varies across countries, and also affects firms behavior. As an example, recently in France and Belgium, several alliances of supermarket chains were inspected by the EU competition authorities. The EU commission was investigating many cases of potential dominant market positions, such as the case of the world's biggest beer brewer AB in Ber, for its anticompetitive behavior in Netherlands and France. Moreover, food prices increased significantly after the 2007 financial crisis and commodity prices have shown increased volatility since then. Better understanding of price formation and transmission has become a crucial issue.

The acts of producing and consuming food can come with externalities to health and the environment. In the health policy realm, for example, governments throughout the EU have contemplated taxes on "sin goods" like sugar and fat, with Ireland and UK most recently gaining approval from the EU to levy a tax on sugary drinks. In environmental policy, pressing concerns about damage from global warming, which are exacerbated by the production of food products with CO_2 intensive production processes like beef, raise concerns of whether taxing such goods might help the environment without greatly reducing human welfare. All of this is to say that having a good understanding of food markets and their workings is of critical importance. Yet food markets can also be complex, with many stakeholders, intermediate inputs, and other moving parts. Hence, exploring the economic underpinnings of these food markets using structural models can often be a useful exercise.

The objective of this paper is to review structural empirical methods that allow tackling relevant policy questions for food markets such as price formation, competition issues, or public policies such as nutritional, environmental, or regulatory policies that could affect food markets. The first key issue of those analyses is the understanding of consumer behavior for food products. How do consumers value food products and their characteristics? What drives the consumer substitution patterns? The second key issue is to determine firms behavior in food and drink markets along the supply distribution chain. Two empirical approaches exist: the reduced form approach and the structural approach resulting from the NEIO (New Empirical Industrial Organization) literature. This second approach solves the problem of reduced-form models that do not explain the link between economic policy decisions and agents' expectations. Indeed, these models estimate the impact of economic policy measures "all things being equal", in particular leaving unchanged the expectations of agents. Structural econometrics can therefore take into account its strategic expectations. It is based on the one hand on the economic modeling of the agents to describe the behavior of the actors of the agri-food chain and on the other hand on statistical hypotheses to infer what one does not observe. The data available on the market do not directly describe consumer substitution patterns or levels of competition upstream or downstream of a value chain. There is also no public data on contracts between the actors or on their margins. Modeling and inference make it possible to overcome this lack of observations. This survey will focus on a two-step approach.[1] The basic idea of two-stage structural modeling can be described following the

[1]Another approach exists where supply and demand are modeled simultaneously. These simultaneous ap-

insight given by Rosse (1970) in the case of the analysis of the market power of a monopoly. Typically, the researcher only has demand data such as prices, quantities, and values of observable product characteristics. A demand model can then be specified, and parameters estimated. On the other hand, the researcher has little information on the production costs of the products and is therefore not able to calculate price-cost margins within a sector. To circumvent this problem, Rosse shows that the marginal cost can be deduced from the first-order condition of the problem of profit maximization of the firm and depends only on the data of prices, quantities, observable variables as well as the estimates of the demand. This intuition can be transposed to the case of agri-food chains composed mainly of oligopolies. The specification of the demand function as well as the assumptions made on the strategic relationships between the different actors in agri-food chains then determine the flexibility of the structural model. In an oligopolistic market where products are differentiated and preferences of heterogeneous consumers, modeling of household purchasing behavior affects price elasticities of demand, that is the percentage change in demand when the price of a product varies, and, therefore, the estimation of the degree of competition within the industry. To study consumer behavior, different approaches are used in the literature: classical demand system models as AIDS, Multilevel Demand System or EASI models, and random utility approaches as Logit, nested Logit, random coefficient Logit, and random coefficient nested Logit models. We will present advantages and drawbacks of those demand models. Regarding the modeling of competitive behavior of the different actors of the food chain, we need to consider vertical relationships between manufacturers and retailers as the latter are

proaches require restrictive statistical assumptions about the relationships between the error terms of the non-necessary supply and demand equations in the two-step approach we present here.

powerful intermediaries that can not be neglected. We will present two different approaches to model vertical interactions: take-it or leave-it offers with linear and non-linear contracts and bargaining game models.

The paper is organized as follows. In section 2 we review different ways empirical researchers have approached the issue of demand estimation in the applied contexts that we typically confront as economists doing applied policy analysis and empirical structural food industrial organization research. Within that context we provide an overview of the state of research and summarize the core problems researchers face when estimating demand. Section 3 turns to the structural supply models implemented in applied research to date and section 4 reviews the main topics in policy analysis using the demand and supply models. Finally, section 5 discusses implications of the main literature findings and contains closing suggestions for future research.

2 Structural Demand Models

Many questions in applied policy analysis require an understanding of how consumers choose among various goods and services as a function of market and individual characteristics. First, properly estimating a demand system in its own right is an objective of interest. Second, demand systems (and their underlying parameters) are often used as providing the "ingredients" to compute consumer welfare from a policy change in a partial equilibrium setting.

The empirical setting of a multiple goods demand specification is a simultaneous equation demand system given by:

(1) $$\log(q_j) = \alpha p_j + \beta p_k + \gamma x_j + \epsilon_j$$

Contained in the unobservable (ϵ_j) are demand shifters that are not in the set of regressors. Prices are endogenous and this, at the very least, calls for a very demanding Instrumental variable strategy, which we will turn to in future sections. Also, as the number of products increases, the number of parameters to be estimated will get very large. We address ways to deal with dimensionality problem next.

The way to reduce the dimensionality of the estimation problem is to put more structure on the choice problem being faced by consumers. In the literature there are two main approaches to tackling such a problem. The first is to model consumers' choices in the actual product space as classical demand systems as well as using specific forms of the underlying utility functions that generate empirically convenient properties, such as the Exact Affine Stone Index (EASI) Marshallian Demand System (Lewbel and Pendakur, 2009); the Almost Ideal Demand System (Deaton and Muellbauer, 1980a) or short AIDS, which is EASI's most used special case; Multilevel demand system (Hausman, 1997); and the Quadratic AIDS model (Banks et al, 1997), or short Q-AIDS, which is more general than AIDS as it allows for quadratic income effects. The second approach is to project products into a smaller dimensional characteristics space and estimate demand that results from a random utility model foundation (McFadden, 1974). We will start with product space approaches and then turn to the attribute space approaches to modeling demand in food markets.

2.1 Classical Demand Systems

2.1.1 Almost Ideal Demand Systems and Multilevel Demand System

The Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980a) is the state of the art for product space approaches. AIDS models were the dominant choice for applied work. In particular, the AIDS model can fit the data much better than more structural models in some applications, and show you just how far you can get with a more "reduced-form" model. These are also used in settings when researchers need to have a more complicated supply side, like for example in a dynamic entry game to ease computational burdens and capture some of the reality of the data generating process. The main disadvantage with AIDS approaches, is that when anything changes in the model (more consumers, adding new products, imperfect availability in some markets), it is difficult to modify the AIDS approach to account for this type of problem.

The usual empirical approach is to use a model of multi-level budgeting. The idea is to impose something akin to a "utility tree" (Hausman, 1997) where consumers first choose to allocate expenditures to an upper level product grouping. Within each product grouping, consumers allocate expenditures among the products therein. When allocating expenditures within a group it is assumed that the division of expenditure within one group is independent of that within any other group. That is, the effect of a price change for a good in another group is only felt via the change in expenditures at the group level. If the expenditure on a group does not change (even if the division of expenditures within it does) then there will be no effect on goods outside that group. To be able to allocate expenditures across groups you have to be able to come up with a price index which can be calculated without knowing what is chosen within the group. These two requirements lead to restrictive utility specifications, the most commonly used being the Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980a).[2]

Starting at the within-group level, assume the expenditure functions for utility u and price vector p look like:

$$\log(e(u,p) = (1-u)\log(a(p)) + u\log(b(p)) \tag{2}$$

where it is assumed that $\log(a(p)) = \alpha_0 + \sum_k \alpha_k \log(p_k) + \frac{1}{2}\sum_k \sum_j \gamma^*_{kj} \log(p_k)\log(p_j)$ and that $\log(b(p)) = \log(a(p)) + \beta_0 \prod_k p_k^{\beta_k}$.

Using Shephard's Lemma, we get shares of expenditure within groups denoted as wg_i and defined as

$$wg_i = \frac{\partial \log(e(u,p))}{\partial \log p_i} = \alpha_i + \sum_j \gamma_{ij} \log(p_j) + \beta_i \log(\frac{x}{P}) \tag{3}$$

where x is total expenditure on the group, $\gamma_{ij} = \frac{1}{2}(\gamma^*_{ij} + \gamma^*_{ji})$, and P is a price index for the group. The price index that "deflates" income is specified as a linear approximation (as in Stone 1954) used by most of the empirical literature, given by $\log(P) = \sum_k w_k \log(p_k)$.

For the allocation of expenditures across groups, we treat the groups as individual goods, with prices being the price indexes for each group. Once again, this depends on the initial choice of groupings. The next step is to calculate expenditure share w_i of each good i using prices p_i, quantities q_i, and total expenditure $x = \sum_k p_k q_k$. Then, we compute the Stone

[2]See Deaton and Muellbauer (1980a), which goes through all the micro-foundations.

price index: $\log P = \sum_k w_k \log(p_k)$, and use IV to estimate the equation:

$$w_i = \alpha_i + \sum_k \gamma_{ik} \log(p_k) + \beta_i \log\left(\frac{x}{P}\right) + \xi_i \tag{4}$$

where ξ_i is the error term. Ultimately we recover $J+2$ parameters $(\alpha_i, \gamma_{i1}, ...\gamma_{iJ}, \beta_i)$, where economic theory imposes that the parameters satisfy several restrictions, namely: homogeneity, symmetry and adding-up.[3]

2.1.2 Exact Affine Stone Index (EASI) Marshallian Demand

Like the AIDS, the Exact Affine Stone Index (EASI) demand system features budget shares that are linear in parameters given real expenditures. However, unlike the AIDS, EASI demands' Engel curves can have any shape over real expenditures, and below we will use a fifth order polynomial, unlike the AIDS, which are linear in income (Lewbel and Penkadur, 2008). It can be thus said that AIDS is a special case of EASI.

The EASI demand system has the following equation for the budget share w_j of each good j:

$$w_j = \sum_{r=0}^{5} b_{rj} y^r + \sum_{d=1}^{Z} (C_{dj} z_d + D_{dj} z_d y) + \sum_{d=1}^{Z} \sum_{k=1}^{J} A_{dkj} z_d p_k + \sum_{k=1}^{J} B_{kj} p_k y + \epsilon_j, \tag{5}$$

where y is a measure of total expenditures entering as a fifth order polynomial. The y is either approximated by the Stone index deflated log nominal expenditure x, namely $y = x - \sum_{j=1}^{J} p_j w_j$ or given by $y = (D, p, x, z) = \frac{x - \sum_{j=1}^{J} p_j w_j + \sum_{d=1}^{Z} z_d p' A_d p/2}{1 - p' B p/2}$. Here, p_k and the log

[3]Note that the Q-AIDS model (see Banks, et al, 1997 for the details) is more general, as it allows for quadratic income effects.

of prices of each good k, the Z demographic characteristics z_d, the interaction terms of the forms $p_k y$, $z_d y$, and $z_d p_k$, and (A, B, C, D) are the parameters to be estimated, subject to homogeneity, symmetry and adding-up demand theory restrictions on the parameters.

2.1.3 Estimation methods

In the literature there is a broad discussion of whether the demand restrictions imposed of homogeneity, symmetry, and adding up hold when estimating models using aggregate choices, namely aggregate expenditures, rather than individual level expenditures (see e.g., Deaton and Muehllbauer, 1980b; and Christiansen, Jorgenson and Lau, 1975, that rejected symmetry and homogeneity in aggregated datasets). Hence, in practice, empirical research has used data on prices, quantities and expenditure, imposing only the adding-up condition. In addition, to estimate the several classical demand systems described above we must deal with the dimensionality and price endogeneity issues.

To reduce the dimensionality issues, demand theory restrictions are imposed on the demand parameters to be estimated. As a consequence, substitution patterns are driven by functional forms. For instance, if we were to look at substitution across segments, the AIDS model restricts substitution patterns to be the same between any two products in different segments. This is not a general assumption and is a restriction that can affect research, if, for policy analysis, we need demand that relies on having estimated elasticity patterns as revealed by consumer purchases and not as dictated by functional forms of the demand model. In addition, the major issue for the AIDS model is the high dimensionality of parameters to be estimated. Researchers then need to do some serious aggregating of products to get rid of this problem. Aggregation into higher level product groupings is therefore a solution followed by most of the literature.

Finally recall that, as usual, price is likely to be correlated with the unobservable. Given panel data on prices, quantities and incomes (expenditures) by brands over time in several markets, Hausman, Leonard, and Zona (1994) propose using the prices in one market to instrument for prices in another market. This works under the assumption that the pricing rule looks like

$$\log(p_{jnt}) = \delta_j \log(c_{jt}) + \alpha_{jn} + \omega_{jnt}$$

where p_{jnt} is the price of good j in market n at time t, c_{jt} represents nation-wide product costs at time t, α_{jn} are market specific shifters which reflect transportation costs or local wage differentials, and ω_{jnt} is a mean zero stochastic disturbance (e.g., local sales promotions).

Here they are claiming that market demand shocks ω_{jnt} are uncorrelated across markets. This allows the use of prices in other markets for the same product in the same time period as instruments (if you have a market fixed effect). Often these are referred to as *Hausman instruments*. This strategy has been criticized for ignoring the phenomena of nation-wide advertising campaigns, suggesting that market demand shocks may be correlated under global markets. Therefore, the validity of "Hausman instruments" is now often questioned and researchers use cost side data on c_{jt} as shifters to instrument for price instead of making assumptions on ω_{jnt} across markets. Still, Hausman instruments have been used in different ways in several different studies. Often people use factor price instruments, such as wages, or the price of raw inputs, as variables that shift marginal costs (and hence prices), but don't affect the demand unobservables (ϵ). Instruments can also be used if there is a large price change in one period for some external reason (like a strategic shift in all the companies' pricing decisions). Then the instrument is just an indicator for the pricing shift having occurred or not.

2.2 Random Utility Models

To model consumer preferences, we can also consider the characteristic space approach. The random utility model considers products as bundles of characteristics, and defines consumer preferences over characteristics, letting each consumer choose a bundle which maximizes their utility. The random utility model restricts the consumer to choosing only one bundle. Multiple purchases are easy to incorporate conceptually but incur a big computational cost and require more detailed data than is usually available. [4] These models use either consumer level data over markets, or aggregate sales data of choices of products over markets.

Let the utility for good j, where $j = 0, 1, 2...J$, of the individual i be given by:

$$U_{ij} = U(x_j, p_j, v_i, \theta). \tag{6}$$

Good 0 is generally referred to as the *outside option* or outside good. It represents the option chosen when none of the observed goods are chosen. A maintained assumption is that

[4] Working on elegant ways around this problem is an open area for research. One approach, presented in Dubé (2004), is multiple-discrete choice, where, recognizing that the time of purchase of a good is often not the actual time of consumption, consumers are modeled as purchasing bundles of goods in anticipation of a stream of distinct consumption occasions before their next trip to the supermarket. Another potential approach to deal with multiple purchases is with multiple discrete-continuous choice models, which allows consumers to choose continuous quantities of multiple goods. Kim, Allenby, and Rossi (2002) and Richards, Gomez, and Pofahl (2011) both do this using the multiple discrete-continuous extreme value model of Bhat (2005), while Song and Chintagunta (2006) applies a "shopping basket" model where consumers choose a utility-maximizing bundle of goods.

the pricing of the outside good is set exogenously. J is the number of goods in the industry at hand. x_j are non-price characteristics of good j, p_j is the price, v_i are characteristics of the consumer i, and θ are the parameters of the model. [5]

Consumer i chooses good j when $U_{ij} > U_{ik}, \forall k \neq j$. This means that the set of consumers that choose good j is given by:

$$A_j(\theta) = [v | U_{ij} > U_{ik}, \forall k] \tag{7}$$

Given a distribution over the v's, $f(v)$, we can recover the share of good j as:

$$s_j(x, p|\theta) = \int_{v \epsilon A_j(\theta)} f(dv) \tag{8}$$

If we let the market size be M then the total demand is $q_j(x, p|\theta) = M s_j(x, p|\theta)$.

Recall that ordinal rankings of choices are invariant to affine transformations of the underlying utility function. More specifically, choices are invariant to multiplication of U by a positive number and the addition of any constant. This also implies that in modeling utility we need to make some normalizations, by bolting down a basis to measure things against, and allowing us the ability to interpret our coefficients and do estimation. Traditionally, we normalize the mean utility of the outside good to zero.

Random utility models allow researchers to address horizontal and vertical product dif-

[5] In these assumptions, we are implicitly supposing that the characteristics of given products are fixed and do not vary over different consumers. This seems sensible, but could actually be an issue if, for example, choice sets of consumers are different (due, for instance, to differences in transportation means that would determine the set of stores they could visit) in ways the researcher cannot observe.

ferentiation. Horizontally differentiated means that, setting aside price, people disagree over which product is best. In the characteristics space setting this means that with horizontal differentiation, people don't all agree that more of an attribute is better. Vertically differentiated means that, price aside, everyone agrees on which good is best, they just differ in how much they value additional quality.

In the following subsections, we will describe three types of random utility models, that make slightly different assumptions about the consumers' utility functions. First, we discuss the Logit model of McFadden (1974), where we assume that consumers' preferences for a product are uncorrelated across all products. Next, we discuss the nested Logit model, which relaxes these strong preference assumptions, and instead allows consumers to have correlated preferences over products in the same subgroup (essentially allowing products in a subgroup to be substitutes). Finally, we discuss the random coefficients Logit model (RCLM), which allows consumers to have idiosyncratic preferences over product attributes that, unlike the previous two models, are not solely based on observable characteristics.

2.2.1 Logit

The Logit model[6] assumes that everyone has the same taste for quality but have different idiosyncratic taste for the product. Utility is given by:

$$U_{ij} = \delta_j + \epsilon_{ij}, \tag{9}$$

where ϵ_{ij} is distributed iid extreme value type I, with $F(\epsilon) = \mathrm{e}^{-\mathrm{e}^{-\epsilon}}$.

[6] See McFadden 1974 for details on the construction.

The probability that good j is chosen by consumer i is given by:

$$\text{Pr}_i(\text{Choice}_j) = \text{Pr}(U_{ji} > U_{hi}) = \text{Pr}(\delta_j + \epsilon_{ji} > \delta_h + \epsilon_{hi}), \forall h \neq j, \forall i. \tag{10}$$

This allows for the aggregate shares, or the probability that good j is chosen across all consumers, to have an analytical form, obtained as:

$$\text{Pr}(\text{Choice}_j) = \frac{\exp(\delta_j)}{\sum_{n=0}^{J} \exp(\delta_n)} = \frac{\exp(\delta_j)}{1 + \sum_{n=1}^{J} \exp(\delta_n)}, \tag{11}$$

where the last simplification in equation (11) results from the fact that $\delta_0 = 0$. That is, the normalized mean utility for the outside option is zero.

This ease in aggregation comes at a cost. The embedded assumption on the distribution on tastes creates more structure than we would like on the aggregate substitution matrix, as well as independence of irrelevant alternatives (IIA). The problem is that the ratio of choice probabilities of consumer i between two options j and k does not depend on the characteristics and or utilities of any other product, i.e. $\frac{S_{ij}}{S_{ik}} = \frac{\exp(\delta_{ij})}{\exp(\delta_{ik})}$ and cross-price elasticities are restricted to satisfy $\frac{\partial s_j}{\partial p_k} = \frac{\partial s_l}{\partial p_k} \; \forall \; l \neq j$.

If we have data with choices of individuals i, and with respondent-specific choice information and respondents' demographics, these data enable us to consider and estimate a specification of heterogeneous preferences. The mean utility δ_{ij} is specified as:

$$\delta_{ij} = \beta_0 x_j + \beta_1 D_i x_j, \tag{12}$$

where x_j are attributes of product j, β_0 is the mean taste for attributes and β_1 is the specific taste according to the respondent's observed demographics D_i. This structure allows for the fact that different decision makers may have different preferences.

2.2.2 Nested Logit

The nested Logit model relaxes the IIA property of the simple Logit model, and allows consumers to have correlated preferences for products that belong to the same subgroup or group in a nest. As in the AIDS Model, we need to make some "ex-ante" classification of goods into different segments, so each good $j \in S(j)$. The goods are divided into nests, and we allow for a degree of independence in unobserved components within each nest. For different goods in different nests, the relative choice probabilities are now dependent on attributes of other alternatives in the two nests (McFadden, 1978, Train, 2003).

Let the set of products be partitioned into K non-overlapping nests denoted as $N_1, N_2, ..., N_K$. The utility that person i obtains from alternative j in nest N_k is denoted, as usual, as $U_{ij} = V_{ij} + \varepsilon_{ij}$, where V_{ij} is observed by the researcher, and ε_{ij} is a random variable where the vector $\varepsilon_i = (\varepsilon_{i1}, ..., \varepsilon_{iJ})$ has a Generalized Extreme Value distribution, which is a generalization of the distribution that gives rise to the Logit, allowing for the ε's to be correlated within a nest. We define a parameter λ_k which measures the degree of independence in unobserved utility among the alternatives in nest k where, if $\lambda_k = 1$ for all nests k, we obtain the Logit model.

Given that consumers choose the product that maximizes random utility, the choice probabilities for each product j in nest N_k are given by

$$\mathrm{Pr}_{ij} = \frac{\exp(V_{ij}/\lambda_k)(\sum_{j \in N_k} \exp(\delta_{ij}/\lambda_k))^{\lambda_k - 1}}{\sum_{l=1}^{K}(\sum_{m \in N_l} \exp(\delta_{im}/\lambda_l))^{\lambda_l}}. \tag{13}$$

We can get a simpler expression for the probability if we rewrite $V_{ij} = W_{ik} + Y_{ij}$ for a good in nest N_k, where W_{ik} depends only on variables that describe nest k that differ over nests but not over alternatives within each nest. In addition, in Y_{ij} we have variables that describe alternative j and vary over alternatives within nest k. With this decomposition of utility, the nested Logit probability can be written as the product of two standard Logit probabilities. Let the probability of choosing alternative $j \in N_k$ be expressed as the product of a conditional probability and a marginal probability, namely $\Pr_{ij} = \Pr_{ij|N_k} \Pr_{iN_k}$, where $\Pr_{iN_k}$ is the probability that an alternative within nest N_k is chosen and $\Pr_{ij|N_k}$ is the probability that the alternative j is chosen given that an alternative in N_k is chosen.

For the nested Logit we get:

$$\Pr_{iN_k} = \frac{\exp(W_{ik} + \lambda_k I_{ik})}{\sum_{l=1}^{K} \exp(W_{il} + \lambda_l I_{il})}, \tag{14}$$

$$\Pr_{ij|N_k} = \frac{\exp(Y_{ij}/\lambda_k)}{\sum_{m \in N_k} \exp(Y_{im}/\lambda_k)} \tag{15}$$

where the so called inclusive value I_{ik} is defined as:

$$I_{ik} = \log \sum_{j \in N_k} \exp(Y_{ij}/\lambda_k) \tag{16}$$

This means that the nested Logit probabilities are ultimately given by:

$$\Pr_{ij} = \frac{\exp(W_{ik} + \lambda_k I_{ik})}{\sum_{l=1}^{K} \exp(W_{il} + \lambda_l I_{il})} \frac{\exp(Y_{ij}/\lambda_k)}{\sum_{m \in N_k} \exp(Y_{im}/\lambda_k)}. \tag{17}$$

2.2.3 Random Coefficient Logit Models

Even if nested Logit model relaxes the IIA assumption of the Logit model, this model is not totally flexible as the IIA assumption holds for two alternatives within the same group. Another stream of papers estimate flexible specifications of discrete choice structural revealed preference models of consumer demand (McFadden, 1974; and McFadden and Train, 2000; Train, 2003) via a random coefficient. Here we allow individual decision makers to have different preferences not just due to observable demographics D_i, which sometimes may not capture all the heterogeneity, but rather allow a more general unobserved heterogeneity structure. Then we define the coefficients β_i to vary according to $\beta_i = \beta_0 + \sigma_\beta v_i^\beta$, where v_i follows a parametric distribution [7] capturing any heterogeneity and the price parameter is similarly defined as $\alpha_i = \alpha_0 + \sigma_\alpha v_i^\alpha$. If there is no heterogeneity in individual preferences relative to the average, then both σ_α and σ_β will be zero and the model will correspond to a Logit model. If, however, there is heterogeneity in preferences relative to the average, then σ_α and σ_β are different from zero.

The utility is defined as:

$$U_{ij} = X_j\beta_i + \alpha_i \text{Price}_j + \xi_j + \varepsilon_{ij}. \tag{18}$$

If ε_{ji} are assumed to be independently, identically extreme value distributed (type I extreme value distribution) and consumers choose one unit of product j among all the possible products available at a certain time that maximizes their indirect utility, the following closed

[7] Train (2016) develops a semi-parametric form of the distribution of preferences.

form solution can be derived for the probability that the consumer i chooses the product j conditional on $v_i = (v_i^\beta, v_i^\alpha)$ is:

$$\mathrm{Pr}_{ji} = \frac{\exp(X_j\beta_i + \alpha_i \mathrm{Price}_j)}{\sum_{k=0}^{J} \exp(X_k\beta_i + \alpha_i \mathrm{Price}_k)} \tag{19}$$

This offers flexibility in incorporating consumer heterogeneity with regard to product attributes X. To recover how D_i affects the departure from mean valuations, we then project estimated β_i on observed demographics D_i in a second step, as described below.

In a third heterogeneity specification, we define the coefficients α_i and β_i to be combination of the two previous heterogeneity specifications, as $\alpha_i = \alpha_0 + \alpha_1 D_i + \sigma_\alpha v_i^\alpha$ and $\beta_i = \beta_0 + \beta_1 D_i + \sigma_\beta v_i^\beta$, where v_i^β and v_i^α are normal random variables capturing any random heterogeneity and D_i are observed consumer characteristics affecting heterogeneity. If, however, there is heterogeneity in preferences due to demographics relative to the average, then α_1 and β_1 are different from zero, and if there is additional random heterogeneity, then σ_α and σ_β are different from zero as well.

2.2.4 Estimation methods

Micro data

If we have micro level data, we see each consumer making decisions for T different choice occasions. By defining the distribution of the $\theta = (\alpha, \beta)$ parameters in general form as $f(\theta|\alpha_0, \alpha_1, \alpha_2, \beta_0, \beta_1, \beta_2)$ and $\alpha = \alpha_0$ for all consumers, then the probability of individual i

making a sequence of choices among the N alternatives $(j = 0, ...N)$ is given as:

$$(20) \qquad S_i = \int \prod_{t=1}^{T} \prod_{j=0}^{J} \left[\frac{\exp(X_{ijt}\beta_i + \alpha_i \text{Price}_{jt})}{1 + \sum_{k=1}^{J} \exp(X_{ikt}\beta_i + \alpha_i \text{Price}_{kt})} \right]^{Y_{ijt}} f(\theta|\alpha_0, \alpha_1, \sigma_\alpha, \beta_0, \beta_1, \sigma_\beta) d\theta,$$

where $Y_{ijt} = 1$ if the respondent i chooses alternative j for situation t and 0 otherwise. Given a total of I respondents, the parameters $(\alpha = \alpha_0, \alpha_1, \sigma_\alpha, \beta_0, \beta_1, \sigma_\beta)$ are estimated by maximizing the simulated log-likelihood function:

$$(21) \qquad SLL = \sum_{i=1}^{I} \log \left(\frac{1}{R} \sum_{r=1}^{R} \prod_{t=1}^{T} \prod_{j=0}^{J} \left[\frac{\exp(X_{ijt}\beta_i^{[r]} + \alpha_i^{[r]} \text{Price}_{jt})}{1 + \sum_{k=1}^{J} \exp(X_{ikt}\beta_i^{[r]} + \alpha_i^{[r]} \text{Price}_{kt})} \right]^{Y_{ijt}} \right),$$

where $\alpha_i^{[r]}, \beta_i^{[r]}$ are the r-th draw for respondent i from the distribution of α, β. [8]

Let $\theta = (\alpha, \beta)$. To estimate θ_i we proceed as follows. The expected value of θ, conditional on a given response Y_i of individual i and a set of alternatives characterized by X_i for product t, is given by:

$$(22) \quad E[\theta|Y_i, Xi] = \frac{\int \theta \prod_{t=1}^{T} \prod_{j=0}^{J} \left[\frac{\exp(X_{ijt}\beta + \alpha \text{Price}_{jt})}{1 + \sum_{k=1}^{J} \exp(X_{ikt}\beta + \alpha \text{Price}_{kt})} \right]^{Y_{ijt}} f(\theta|\alpha_0, \alpha_1, \sigma_\alpha, \beta_0, \beta_1, \sigma_\beta) d\theta}{\int \prod_{t=1}^{T} \prod_{j=0}^{J} \left[\frac{\exp(X_{ijt}\beta + \alpha \text{Price}_{jt})}{1 + \sum_{k=1}^{J} \exp(X_{ikt}\beta + \alpha \text{Price}_{kt})} \right]^{Y_{ijt}} f(\theta|\alpha_0, \alpha_1, \sigma_\alpha, \beta_0, \beta_1, \sigma_\beta) d\theta},$$

Equation (21) can be thought as the conditional average of the coefficient for the sub-group of individuals who face the same alternatives and make the same choices. For each individual i we estimate a certain attribute marginal utility θ_i, following Revelt and Train (2000) , by

[8] Normal distribution is common assumption. Log normal distribution can be used to impose positivity of parameters.

simulation according to the following:

$$\hat{\theta}_i = \frac{\frac{1}{R}\sum_{r=1}^{R} \beta_i^{[r]} \prod_{t=1}^{T} \prod_{j=0}^{J} \left[\frac{\exp(X_{ijt}\beta_i^{[r]} + \alpha_i^{[r]} \text{Price}_{jt})}{1+\sum_{k=1}^{J} \exp(X_{ikt}\beta_i^{[r]} + \alpha_i^{[r]} \text{Price}_{kt})} \right]^{Y_{ijt}}}{\frac{1}{R}\sum_{r=1}^{R} \prod_{t=1}^{T} \prod_{j=0}^{J} \left[\frac{\exp(X_{ijt}\beta_i^{[r]} + \alpha_i^{[r]} \text{Price}_{jt})}{1+\sum_{k=1}^{J} \exp(X_{ikt}\beta_i^{[r]} + \alpha_i^{[r]} \text{Price}_{kt})} \right]^{Y_{ijt}}}, \tag{23}$$

where $\theta[r]$ is the r-th draw for individual i from the estimated i's distribution of θ.

If prices are exogenous, the demand parameters are estimated by simulated Maximum Likelihood function defined in (21). However, the price exogeneity assumption can not be hold when omitted product characteristics affect both demand and prices. Omitted characteristics could be unobserved product attributes of goods or unobserved marketing efforts as advertising, sales promotions, shelves position. Then, the method used follows simulated likelihood function control approaches as in Kuksov and Villas-Boas (2008), and Petrin and Train (2010). This method consists in regressing prices on the exogenous demand variables and instrumental variables, and capturing, through the error term of the price equation, all unobserved product characteristics that affect prices. This error term is then included in the deterministic part of the utility to avoid any correlation between prices and the remaining part of the demand error term. The choice of instrumental variables is crucial. They should be price shifters uncorrelated with demand. The ideal instrumental variables should capture the cost of producing and distributing the goods, and inputs costs are good candidates, as are attributes of other products (as in Berry et al., 1995) or Hausman style instruments such as average prices in other markets, assuming no national common demand unobservables.

Aggregate Sales Data

The random coefficient Logit model using aggregate data, as in BLP (1995) and Nevo (2000a), is probably the most commonly used demand model in the empirical food industry literature. This model includes differentiated goods, and the data are typically shares, prices, and characteristics in different markets. All the standard problems, such as being endogenous and wider issues of identification, will continue to be a problem here. Estimation of this model uses variation in prices due to exogenous factors to identify demand.

As in Nevo (2000a), we rewrite the utility of consumer i for product j as:

$$U_{ijt} = \delta_{jt}(p_{jt}, X_{jt}, \xi_{jt}; \alpha_0, \beta_0) + \mu_{ijt}(p_{jt}, v_i; \sigma_\beta, \sigma_\alpha), \tag{24}$$

where $\delta_{jt} = \alpha_0 p_{jt} + \beta_0 X_{jt} + \xi_{jt}$ is the mean utility, while $\mu_{ijt} = \sigma_\beta X_{jt} v_i^\beta + \sigma_\alpha p_{jt} v_i^\alpha + \varepsilon_{ijt}$ is the deviation from the mean utility that allows for consumer heterogeneity in price response.

Let the distribution of μ_{ijt} across consumers be denoted as $F(\mu)$. The aggregate share S_{jt} of product j in market t across all consumers is obtained by integrating the consumer level probabilities:

$$S_{jt} = \int \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{n=1}^{J} \exp(\delta_{nt} + \mu_{int})} dF(\mu). \tag{25}$$

This aggregate demand system not only accounts for consumer heterogeneity, but also provides more flexible aggregate substitution patterns than the homogeneous Logit model. If price is correlated with demand unobservables ξ then the coefficients are not consistently estimated by OLS.

Using aggregate data on market shares by product over markets, to estimate the random parameters Logit demand model the literature uses the GMM-estimator proposed by Berry, Levinsohn, and Pakes (1995) and Nevo (2001a). When estimating demand, the goal is to derive parameter estimates that produce product market shares close to the observed shares. As shown, this procedure is non-linear in the demand parameters to be estimated.

Berry (1994) constructs a demand side equation that is linear in the parameters to be estimated. This follows from equating the estimated product market shares.[9] Once this inversion has been made, given starting values of the random coefficients σ_β and σ_α, one obtains equation (26) which is linear in the linear parameters, and then instrumental variable methods can be applied directly.

$$\delta_{jt}(\sigma_\beta, \sigma_\alpha) = \alpha p_{jt} + X_{jt}\beta + \xi_{jt}.. \tag{26}$$

Finally, the random coefficient parameters σ_β and σ_α are obtained by feasible Simulated Method of Moments (SMOM) following Nevo's (2000a) estimation algorithm using equation (26).

3 Structural Supply Models of Food Markets

Empirical work in food supply chains either assumes manufacturers sell directly to consumers, or that the intermediate distributor or retailer remains passive (as in Baker and Bresnahan,

[9] The product market share in is approximated by the Logit smoothed accept-reject simulator. Solving for the mean utility (as in Berry 1994) has to be done numerically (see Berry, Levinsohn, and Pakes, 1995; and Nevo, 2001a).

1985; Berry and Pakes, 1993; Werden and Froeb, 1994; Nevo, 2000a), or models the supply chain strategic pricing formally. In these models, manufacturing decide in a take it or leave it setting the wholesale prices (Villas-Boas and Zhao, 2005; Villas-Boas 2007a) or non linear contracts (Bonnet and Dubois, 2010) when contracting or instead via bargaining with retailers (Misra and Mohanty, 2006; Draganska et al., 2010; Bonnet, Bouamra-Mechemache, and Richards, 2018), who in turn decide the retail prices that consumers have to pay.

Consider a set of J products sold by R retailers and M manufacturers. Each retailer r sells a subset S^r of products and each manufacturer m sells to retailers a subset G^m of products, such that $J = \sum_{r=1}^{R} S^r = \sum_{m=1}^{M} G^m$.

3.1 Take-it or Leave-it Offers

3.1.1 Linear Contracts

We assume a manufacturer Stackelberg model in which M manufacturers set wholesale prices p^w first, in a Nash-Bertrand manufacturer-level game, and then R retailers (chains) follow, setting retail prices p in a Nash-Bertrand fashion. Let each retailer's r marginal costs for product j be given by c_j^r, and let manufacturers' marginal costs be given by c_j^w.[10]

Solving by backwards induction, we assume that each retail chain r maximizes his profit function defined by:

$$\pi_r = \sum_{j \epsilon S_r} M \left[p_j - p_j^w - c_j^r\right] s_j(p) \quad \text{for} \quad r = 1, ...R, \tag{27}$$

[10]One can also assume that the manufacturers depart from Bertrand Nash, namely, that they maximize joint profits over the set of products they all produce (Villas-Boas,2007).

where M is the size of the market, and s_j is defined, given a potential market, as the market share of product j. The first-order conditions, assuming a pure-strategy Nash equilibrium in retail prices, are:

$$s_j + \sum_{m \epsilon S_r} [p_m - p_m^w - c_m^r] \frac{\partial s_m}{\partial p_j} = 0 \text{ for } j \in S^r \tag{28}$$

Switching to matrix notation, let Δ_r be a matrix with general element $\Delta_r(i, j) = \frac{\partial s_j}{\partial p_i}$, containing consumer demand substitution patterns with respect to changes in the retail prices of all products, and I^r the diagonal ownership matrix of retailer r where element (j, j) takes 1 if j is sold by the retailer r and 0 otherwise. Solving (28) for the price-cost margins for all products in vector notation gives the price-cost margins γ for all the products in the retail chains under Nash-Bertrand pricing:

$$\gamma = p - p^w - c^r = -\sum_{r=1}^{R} [I_r \Delta_r I_r]^{-1} I_r s(p), \tag{29}$$

which is a system of J implicit functions that expresses the J retail prices as functions of the wholesale prices.

Manufacturers choose wholesale prices p^w to maximize their profits given by:

$$\pi_m = \sum_{j \epsilon G_m} [p_j^w - c_j^w] \; s_j(p(p^w)), \tag{30}$$

knowing that retail chains behave according to (29).[11] Solving for the first-order conditions

[11] Note that in this market manufacturers may, if they choose to, set different wholesale prices for the same brand sold to different retailers. In another study, Villas-Boas (2007b) considers the welfare effects from

from the manufacturers' profit-maximization problem, assuming again a pure-strategy Nash equilibrium in wholesale prices and using matrix notation, yields:

$$\Gamma = (p^w - c^w) = -\sum_{f=1}^{M}[I_f \Delta_w I_f]^{-1} I_f s(p), \tag{31}$$

where I_f is a diagonal ownership matrix of manufacturer f with element $(i,j) = 1$, if the manufacturer sells product j, and equal to zero otherwise; Δ_w is a matrix with general element $\Delta_w(i,j) = \frac{\partial s_j}{\partial p_i^w}$ containing changes in demand for all products when wholesale prices change. Given retail mark-up pricing behavior assumed in (29), Δ_w can be decomposed as the impact of retail prices on demand and the impact of wholesale prices on retail prices, to obtain $\Delta_w(i,j) = \sum_{l=1}^{J} \frac{\partial s_j}{\partial p_l}\frac{\partial p_l}{\partial p_i^w}$. In matrix notation, we get $\Delta_w = P_w \Delta_r$ where P_w is the matrix of the derivatives of retail prices with respect to wholesale prices, and can be deduced from the total differentiation of the retailer's first order conditions (28) with respect to wholesale price. Defining $S_p^{p_j}$ the $(J \times J)$ matrix of the second derivatives of the market shares with respect to retail prices whose element (l,k) is $\frac{\partial^2 s_k}{\partial p_j \partial p_l}$, we obtain P_w as[12]:

$$P_w = I_r \Delta_r' I_r \left[\Delta_r I_r + I_r \Delta_r' I_r + (S_p^{p_1} I_r \gamma | ... | S_p^{p_J} I_r \gamma) I_r\right]^{-1}. \tag{32}$$

Under the above model, given the demand parameters θ, the implied price-cost margins for all J products can be calculated as $m^r(\theta)$ for the retailers and $m^w(\theta)$ for the manufacturers.[13]

imposing uniform wholesale pricing restrictions in the coffee market.

[12] We use the notation $(a|b)$ for horizontal concatenation of a and b.

[13] If the profit maximizing retail mark-up, $\gamma(\theta)$ is non-varying with quantity, then the linear pricing model is

3.1.2 Non Linear Contracts

Bonnet and Dubois (2010) consider that manufacturers have all market power and simultaneously propose take-it or leave-it offers of two-part tariff contracts to each retailer. These contracts are public information and involve specifying franchise fees and wholesale prices.

In the case of these two-part tariff contracts, the profit function of retailer r is:

$$\Pi^r = \sum\nolimits_{j \in S_r} [M(p_j - p_j^w - c_j^r) s_j(p) - F_j] \tag{33}$$

where F_j is the franchise fee paid by the retailer for selling product j.

Manufacturers set their wholesale prices p_k^w and the franchise fees F_k in order to maximize profits equal to:

$$\Pi^m = \sum\nolimits_{k \in G_m} [M(p_k^w - c_k^w) s_k(p) + F_k] \tag{34}$$

for firm m, subject to retailers' participation constraints $\Pi^r \geq \overline{\Pi}^r$, for all $r = 1, .., R$, where $\overline{\Pi}^r$ is an exogenous outside option of retailer r. The expressions for the franchise fee F_k of the binding participation constraint can be substituted into the manufacturer's profit (34) to obtain the following profit for firm m:

$$\Pi^m = \sum\nolimits_{k \in G_m} (p_k - c_k^w - c_k) s_k(p) + \sum\nolimits_{k \notin G_m} (p_k - p_k^w - c_k) s_k(p) - \sum\nolimits_{j \notin G_m} F_j \tag{35}$$

---

indistinguishable from a model where retailers charge a constant retail mark-up $\gamma_{constant}$, if $\gamma_{constant} = \gamma(\theta)$. For special cases of demand models where $\frac{\partial \gamma(\theta)}{\partial q} = 0$ this may be true. For general demand models this is not the case.

This shows that each manufacturer fully internalizes the entire margins on his products but internalizes only the retail margins on rivals' products. Note that the additional term $\sum_{j\notin G_m} F_j$ is constant for the manufacturer m and thus maximizing the profits of m is equivalent to maximizing the sum (35) without this term. They then set wholesale prices in the following maximization program:

$$\max_{\{w_k\}\in G_m} \sum_{k\in G_m} (p_k - \mu_k - c_k)s_k(p) + \sum_{k\notin G_m} (p_k - w_k - c_k)s_k(p).$$

Then, the first order conditions are, for all $i \in G_m$:

$$\sum_k \frac{\partial p_k}{\partial w_i} s_k(p) + \sum_{k\in G_m} \left[(p_k - \mu_k - c_k) \sum_j \frac{\partial s_k}{\partial p_j}\frac{\partial p_j}{\partial w_i} \right] + \sum_{k\notin G_m} \left[(p_k - w_k - c_k) \sum_j \frac{\partial s_k}{\partial p_j}\frac{\partial p_j}{\partial w_i} \right] = 0$$

These conditions allow us to estimate the price-cost margins, given demand parameters.

Then the first order conditions become (in matrix notation), for all $i \in G_f$:

$$I_f P_w s(p) + I_f P_w \Delta_r I_f \Gamma + I_f P_w \Delta_r \gamma = 0.$$

This implies that the manufacturer price cost margin is:

$$\Gamma = \sum_{f=1}^{M} \left(I_f P_w \Delta_r I_f\right)^{-1} \left[-I_f P_w s(p) - I_f P_w \Delta_r \gamma\right] \tag{36}$$

which allows for an estimate of the price-cost margins with demand parameters, using (29) to replace γ and (32) for P_w.

Note that different vertical restraints can be used with two-part tariff contracts. Indeed, manufacturers can also choose retail prices in the case where manufacturers use resale price maintenance as in Bonnet and Dubois (2010) or uniform pricing (Bonnet et al, 2013).

3.2 Bargaining models

Among the most recent approaches to modeling vertical food supply chains, Draganska et al (2010), and Bonnet and Bouamra-Mechemache (2016) allow for non-unilateral bargaining power and model the price negotiation between retailers and manufacturers using bargaining in the German coffee market and in the French fluid milk market, respectively. This methodology enables researchers to estimate the bargaining power for each pair of manufacturers and retailers and infer the resulting profit sharing between them, contrary to take-it or leave-it offers that consider that all market power is in the retailers' or manufacturers' hands.

These models specify the vertical channel as a two-tier industry consisting of M upstream firms and R downstream retailers. Retailers' profit functions are given by (27). Retail margins result from the retailers' choice of final prices and the maximization of the retail profit. Retailers are assumed to compete with each other in Bertrand-Nash fashion in the goods market and set prices for each product as in (29).

Then, wholesale price equilibrium results from the negotiation between firms and retailers and negotiation on wholesale prices is modeled as a Nash bargaining game. Each pair of firms and retailers are assumed to secretly and simultaneously contract over the wholesale price of the product j. Moreover, firms and retailers have rational expectations, such that the ultimate equilibrium outcome is anticipated by both parties.[14] The main difficulty

[14] In this case, the wholesale prices are determined independently of possible changes to retail prices. For a

comes from the linkage across negotiations, which raises arduous questions: a key difficulty is identifying what each manufacturer knows about its rivals' contract terms. Indeed, when negotiating, each manufacturer must conjecture the set of terms its rivals have or have not been offered. In equilibrium, this conjecture must be correct but out-of-equilibrium beliefs may be important in determining the bargaining outcome. In the cooperative bargaining approach, this problem is solved by assuming that any bargaining outcome must be bilaterally renegotiation-proof, i.e., no manufacturer-retailer pair can deviate from the bargaining outcome in a way that increases their joint profit, taking as given all other contracts. Bargaining between each retailer-manufacturer pair is assumed to maximize the two players' joint profit, taking as given all other negotiated contracts and that each player earns its disagreement payoff (i.e., what it would earn from the sales of its other products if no agreement on this product is reached) plus a share $\lambda_j \in [0,1]$ (respectively $1-\lambda_j$) of the incremental gain from trade going to the retailer (respectively to the manufacturer). A manufacturer negotiates with a given retailer for each of its products, that each product is negotiated separately with the manufacturer, and that retail prices are not observable when bargaining over the wholesale prices. Then, retail prices are considered as fixed when solving for the bargaining solution.

The equilibrium wholesale price for product j is derived from the bilateral bargaining problem between a firm and a retailer such that each firm and retailer pair maximizes the Nash product $\left[\pi_j^r(p_j^w) - d_j^r\right]^{\lambda_j} \left[\pi_j^m(p_j^w) - d_j^m\right]^{(1-\lambda_j)}$ where $\pi_j^m(p_j^w)$ and $\pi_j^r(p_j^w)$ are, respec-

discussion on the wholesale price negotiation and retailer competition games and the justification of the related assumptions, see Draganska et al.(2010) and the literature mentioned therein. Bonnet, Bouamra-Mechemache, and Richards (2018) relax this assumption.

tively, the profits of the firm and the retailer for product j. They are given by:

(37)
$$\begin{aligned} \pi_j^m(p_j^w) &= \left(p_j - p_j^w - c_j^r\right) M s_j(p) = \gamma_j M s_j\left(p\right) \\ \pi_j^r(p_j^w) &= \left(p_j^w - c_j^w\right) M s_j(p) = \Gamma_j M s_j(p). \end{aligned}$$

The payoffs the manufacturer and the retailer can realize outside of their negotiations are denoted, respectively, d_j^m and d_j^r. The retailer could gain d_j^r if it removes the supplier's product j from its stores but contracts with other suppliers. Similarly, the firm could get profits d_j^m from the sales of its other products as well as from the sales of products to other retailers if the negotiation fails. If the retail prices are fixed during the negotiation process, the disagreement payoffs d_j^m and d_j^r are given by:

(38a)
$$\begin{aligned} d_j^r &= \sum_{k \in R^r - \{j\}} \gamma_k M \Delta s_k^{-j}(p) \\ d_j^m &= \sum_{k \in G^m - \{j\}} \Gamma_k M \Delta s_k^{-j}(p) \end{aligned}$$

where the term $M\Delta s_k^{-j}(p)$ is the change in market shares of product k that occurs when the product j is no longer sold on the market. Those quantities can be derived through the substitution patterns estimated in a random coefficient Logit demand model as follows:

(39)
$$\Delta s_k^{-j}(p) = \int \frac{\exp(\delta_k + \mu_{ik})}{1 + \sum_{l=1\backslash\{j\}}^{J} \exp(\delta_l + \mu_{il})} - \frac{\exp(\delta_{kt} + \mu_{ik})}{1 + \sum_{l=1}^{J} \exp(\delta_l + \mu_{il})} dP_\nu(\nu).$$

Solving the bargaining problem in equation (**??**) leads to the following first-order condition:

$$\lambda_j\left(\pi_j^m - d_j^m\right)\frac{\partial \pi_j^r(p_j^w)}{\partial p_j^w} + (1-\lambda_j)\left(\pi_j^r - d_j^r\right)\frac{\partial \pi_j^m(p_j^w)}{\partial w_j} = 0. \tag{40}$$

Under the assumption that the matrix of prices for final commodities is treated as fixed when the wholesale prices are decided during the bargaining process, we have $\frac{\partial \pi_j^r(w_j)}{\partial w_j} = -Ms_j(p)$ and $\frac{\partial \pi_j^m(w_j)}{\partial w_j} = Ms_j(p)$ from equation (37). Equation (40) can thus be written as $\pi_j^m - d_j^m = \frac{1-\lambda_j}{\lambda_j}(\pi_j^r - d_j^r)$. Using equations (37) and (38a), the following expression can be derived for the bargaining solution:

$$\Gamma_j M s_j(p) - \sum_{k\in S^r-\{j\}} \Gamma_k M \Delta s_k^{-j}(p) = \frac{1-\lambda_j}{\lambda_j}\left[\gamma_j M s_j\,(p) - \sum_{k\in G^f-\{j\}} \gamma_k M \Delta s_k^{-j}(p)\right]. \tag{41}$$

Using equation (41) for all products j, we obtain the matrix of firms' margins:

$$\Gamma = \sum_{f=1}^{M}\sum_{r=1}^{R}\left[\left(I_f S I_f\right)^{-1}\left(\frac{1-\lambda}{\lambda} * \left(I_r S I_r\right)\gamma\right)\right].^{15} \tag{42}$$

where S is the $(J\times J)$ matrix with market shares as diagonal elements and changes in market shares otherwise:

$$S = \begin{bmatrix} s_1 & -\Delta s_2^{-1} & \cdots & -\Delta s_J^{-1} \\ -\Delta s_1^{-2} & s_2 & \cdots & -\Delta s_J^{-2} \\ \vdots & \vdots & \ddots & \vdots \\ -\Delta s_1^{-J} & -\Delta s_2^{-J} & \cdots & s_J \end{bmatrix} \tag{43}$$

[15] The $*$ means an element by element multiplication between the vectors $\frac{1-\lambda}{\lambda}$ and $\left[\left(I_f S I_f\right)^{-1}\frac{1-\lambda}{\lambda}\left(I_r S I_r\right)\gamma\right]$

Equation (42) shows the relationship between the wholesale margin on the one hand and the retail margin on the other hand. This relationship first depends on the disagreement payoffs and thus on the market share changes that are determined by the substitution patterns estimated in the demand model. It also depends on the exogenous parameter λ_j, the relative power of the retailer relative to the firm when bargaining over the wholesale price. The higher λ_j, the lower the share of the joint profit the firm will get from the bargaining.

Adding equations (42) and (29) yields the total margin of the firm/retailer pair over product j:

$$\gamma + \Gamma = \left[\sum_{f=1}^{M} \sum_{r=1}^{R} (I_f S I_f)^{-1} \left(\frac{1-\lambda}{\lambda} (I_r S I_r) \right) + I \right] (I_r S_p I_r)^{-1} I_r s(p) \tag{44}$$

where I is the $(J \times J)$ identity matrix.

Because we do not directly observe firms' marginal production costs as well as retailers' marginal distribution costs, we are not able to determine analytically the bargaining power parameter λ_j. We rather conduct an estimation specifying the overall channel marginal cost C_{jt} for each product j. We follow the following specification for the total marginal cost $C_j = \theta\omega_j + \eta_j$ where ω is a vector of cost shifters and η is a vector of error terms that accounts for unobserved shocks to marginal cost. The final equation to be estimated is thus given by:

$$p = \theta\omega + \left[\sum_{f=1}^{M} \sum_{r=1}^{R} (I_f S I_f)^{-1} \left(\frac{1-\lambda}{\lambda} (I_r S I_r) \right) + I \right] (I_r S_p I_r)^{-1} I_r s(p) + \eta. \tag{45}$$

We are then able to get an estimate of λ for each product. Hence, we can deduce manufacturers' margins from equation (42). Moreover, from the estimates of the cost shifters and the error term of equation (45), we get the estimated total marginal cost, which is the sum of the marginal cost of production and the marginal cost of distribution for each product j.

4 Topics in Policy Analysis

In this section, we will consider how the models presented in the previous sections can be useful for policy analysis in food markets. Hence we briefly review papers that utilize structural demand and/or supply models in order to better understand food policy issues. This review is not necessarily meant to be comprehensive, but rather to brush on important themes in the literature, including vertical relationships in food markets, price formation, competition issues, environmental and nutritional policies. Counterfactual food policy analysis through simulations are used to recover variation in prices, consumption, consumer welfare, and profit of policy scenario.

4.1 Vertical relationships

Data on contracts and relationships between manufacturers and retailers in food markets are generally not publicly available. Structural empirical models allow palliating this lack of information making some assumptions about the nature of competition and vertical contracts. Vuong (1989), Smith (1992), and Rivers and Vuong (2002) propose and refine likelihood ratio, Cox-type, and encompassing tests for a variety of nested and non-nested supply models, providing a foundation for testing different structural models. Bonnet and Dubois (2010) find empirical support in the French bottled water market for non-linear pricing and Resale

Price Maintenance (RPM) practices using a random coefficient Logit demand and testing linear and non-linear pricing contracts. This result is corroborated by Bonnet and Réquillart (2013a) in the French soft drink market, Bonnet et al. (2013) in the German coffee market, Villas-Boas (2007a) in the US yogurt market, and Bonnet and Réquillart (2015) in the fluid milk and dairy dessert sectors in France.

Structural models can also be designed to determine bargaining power using the methods presented in Section 3.2. Draganska et al. (2010) evaluate how bargaining power is split throughout the distribution channel in the German coffee market, finding that bargaining power is not an inherent characteristic of a firm but instead depends on the negotiation partner. Additionally, applications with structural models show that product differentiation like organic labeling (Bonnet and Bouamra-Mechemache 2016 and Richards et al. 2011), product complementarity (Bonnet et al. 2018), and introduction of private labels (Meza and Sudhir 2010) are all important determinants of bargaining power along the distribution channel.

The nature of the shock could also affect the pass-through rates. for example, Bonnet and Villas-Boas, allowing for demand asymmetry responses for consumers in a randon coefficient logit approach, show that cost pass-through is higher from a positive cost shock than a negative one. Another example show that the design of the regulatory tools as taxation policies could affect the transmission of the tax. Bonnet and Réquillart (2013b) and Griffith et al. (2018) show that an excise tax will be overtransmitted to consumer while ad-valorem tax leads to a lower consumer incidence on the Frecnh soft drink market and the UK butter and margarine market respectively.

4.2 Competition Issues

Structural approaches are particularly popular for merger and collusive behavior analysis. Slade (2004), for instance, uses a structural approach to look for evidence of collusive behavior in the UK brewing market and finds no evidence of any sort of coordinated behavior. Merger simulations, a prominent application of the tools discussed above, consist in assuming that both firms form a new entity and recovering new price equilibrium given this new structure. Nevo (2001a) is the classical paper on merger analysis in food markets using random coefficient Logit models while Hausman et al. (1994) is the primary reference for merger analysis in food markets using AIDS. Both papers estimate a demand system and resulting firm profit margins given pre-merger pricing behavior. Extending previous merger approaches, Villas-Boas (2007b) finds empirical evidence that authorities should consider incorporating the role of retailers in upstream merger analysis, especially in the presence of an increasingly consolidated retail food industry. Bonnet and Dubois (2010) also extend merger analysis to cases of use of nonlinear contracts by manufacturers and retailers.

4.3 Price Formation

The price formation literature seeks to understand how input price changes throughout the supply chain are passed through into intermediate and final prices for food items, and specifically explain situations where this pass-through is imperfect, considering explanations such as price rigidities, local non-traded costs or markup adjustment of manufacturers and retailers due to consumer substitution patterns, market structure, and market power in industries.

Goldberg and Hellerstein (2008) develop a structural approach that can be used to identify the determinants of incomplete exchange-rate pass-through (one prominent example of incomplete pass-through), exploring markup, marginal cost, and nominal-rigidity channels. Goldberg and Hellerstein (2013) build on this approach, using a random coefficient model of demand to explore incomplete pass through in the beer market, and highlight the role of local non-traded costs. This result is corroborated in the coffee market by Nakamura and Zerom (2010) in the US and by Bettendorf and Verboven (2000) in the Netherlands. Considering other types of pass-though, Nakamura (2008) finds that price variation observed over time seems to arise from retail-level rather than manufacturer-level demand and supply shocks, when looking at US grocery purchases.

Markup adjustment could be due to the degree of competition. Indeed, Kim and Cotterill (2008) looks at cost pass through of changes in inputs for US processed cheese using a mixed logit framework under different regimes, and find that under collusion, the pass-through rates for all brands are much lower than under Nash-Bertrand competition. Nakamura and Zerom (2010) also show that the elasticity patterns of demand affect the markup adjustment in the imperfect pass-through. Vertical integration structures determine pass-through as well; Hellerstein and Villas-Boas (2010) show that pass-through is positively related to firms' degree of vertical integration across industries. Bonnet et al. (2013) find that retailers are limited in their ability to make mark-up adjustments when faced with resale price maintenance restrictions in a context of non linear pricing contrats, leading to higher pass-through rates.

Heterogeneity in firms, including differences in market size or structure also determine pass-through. Atkeson and Burstein (2008), for instance, consider a model of Cournot competition in which firms do not fully pass through changes in their marginal costs to prices because their optimal markup depends on market share. Auer and Schoenle (2016) show that the firms that react the most to changes in their own costs also react the least to changing prices of competing importers. Firm size can also matter; Berman et al. (2011) find that higher-performance French exporter firms react to a currency depreciation by increasing their mark-up significantly more and by increasing their export volume less than other firms.

4.4 Food Policy

To fight against unhealthy food consumption or the carbon foot print of the food consumption, public authorities need to think about effictive policy tools to change food consumption. Structural approaches provide a suitable ex-ante analysis method to anticipate consumers and firms reaction to food policy.

4.4.1 Nutritional Labeling

Structural demand models are used in the literature to understand nutritional label preferences and measure their effect on consumption.[16] In the milk industry, Dhar and Foltz (2005) use data from US cities and Q-AIDS method to estimate the value of organic and rBST-free labels on milk, and find consumer benefits from organic milk and to a lesser extent from rBST-free milk, suggesting value from the corresponding characteristics labels. Kiesel and Villas-Boas (2007) also look at the premium consumers place on a USDA organic label using a random coefficient Logit framework, and find that the USDA organic seal increases

[16] See Kiesel et al., (2011) for a review of various demand approaches to estimating demand in the presence of nutritional labels.

the probability of purchasing organic milk. Brooks and Lusk (2010) similarly find that consumers are willing to pay large premiums to avoid milk from cloned cows. Allais et al. (2015) evaluate front-of-pack nutrition labels and nutrition taxes in the dessert yogurt and fromage blanc markets. Using a supply model of oligopolistic price competition, they find that both taxes and food labels both are effective in reducing the purchase of fatty foods.

4.4.2 Fiscal policies

An alternate method to nudge consumers to make healthier nutritional choices is to tax "sin goods", such as sugar and fat. Allais et al. (2010) look at the effects of a "fat tax" (special tax on food items high in calories, fat, or sugar) on the nutrients purchased by French households across different income groups estimating a complete AIDS. They find that taxing cheese, butter, cream, sugar-fat products reduces total calories purchased, particularly with respect to saturated fat (but also reduced overall food purchases). Bonnet and Réquillart (2011) show with a random coefficient Logit model that some European CAP reforms could damage the nutritional objectives of public authorities, and specifically that the reduction of the price of raw sugar by 36% would increase the consumption of sugar-sweetened drinks in France.

Given the environmental impact of food consumption, several structural analyses address the issue of environmental taxes on food products. For example, Bonnet et al. (2018) implement a random coefficient Logit model to look at how food consumption behavior could change under environmental taxes on all animal products. They find that high taxes €200 per ton of CO_2 would only lead to a 6% reduction in GHG emissions. Caillavet et al.

(2016) and Edjabou and Smed (2013), using an AIDS demand model, also show that the potential for emission reduction is low in France and Denmark, respectively.

5 Conclusion

This review presented structural demand and supply models useful for food policy analysis. Improvements in methodology have paid particular attention to model consumers' and firms' behavior in food markets, and structural approaches are increasingly used as computational power and the availability of data steadily rise. Structural models become especially powerful as more accurate measurements on cost, contractual terms and profit shares become available, as these data are helpful in dealing with various identification issues posed above. Restrictive assumptions remain in the methodologies described and further improvements are required to understand complex behaviors of firms and consumers.

Firms' strategic behavior should be analyzed while accounting for the complexity of their multiple decisions (e.g. price, quantity sold, quality, and variety), suggesting that structural approaches should be improved in the context of vertical relationships. Given the high concentration of unfair practices in the food market, competition authorities will need more and more precise studies of these practices, such as: as tying contracts, cartels, and exclusive dealings, as well as their implications on food supply and profit sharing. As issues of the revenue of upstream actors (producers or producers' organizations) in the agri-food chain are currently being debated in Europe, policy analyses are required to deal with a three-level interactions between producers, manufacturers and retailers. One of the next challenges of structural approaches is to develop vertical interaction models including the upstream sector.

This would allow understanding the implications of Common Agricultural Policy reforms on producers, concentration moves at the manufacturer or retail levels, or environmental and health public policies.

There are also methodological avenues of future research given that the majority of the work in agri-food organization has focused on a single product type and is static in its core approach. A first step would be to extend consumer choices allowing for multiple good choices in the context of multi-category approaches, as well as incorporating firm level, multi-product, strategic behavior. Allowing for dynamics in demand as well as on the supply side is also another avenue to extend the state of the art in agri-food applied work. In so doing, topics that are inherently dynamic in nature could be addressed by empirical work pertaining to innovation, storage, new product introductions, and predatory pricing and other dynamic anti-competitive strategies. Finally, integrating recent work in the field of behavioral economics, which suggests consumers often do not make "rational" choices in the lens of the standard neoclassical framework, into structural models is a promising area for future research as well.

References

[1] Allais, O., Bertail, P. and Nichle, V. (2010). The effects of a fat tax on French households purchases: a nutritional approach. *American Journal of Agricultural Economics* 92(1): 228-245.

[2] Allais, O., Etilé, F. and Lecocq, S. (2015). Mandatory labels, taxes and market forces: An empirical evaluation of fat policies. *Journal of Health Economics* 43: 27-44.

[3] Atkeson, A. and Burstein, A. (2008). Pricing-to-Market, Trade Costs, and International Relative Prices. *American Economic Review* 98(5): 1998-2031.

[4] Auer, R.A. and Schoenle, R.S. (2016). Market structure and exchange rate pass-through. *Journal of International Economics* 98: 60-77.

[5] Banks, J., R. Blundell, and A. Lewbel (1997): Quadratic Engel Curves and Consumer Demand. *The Review of Economics and Statistics* 79(4):527539.

[6] Baker, J.B. and Bresnahan, T.F. (1985). The gains from merger or collusion in product-differentiated industries. *The Journal of Industrial Economics* 1985: 427-444.

[7] Berman, N., Martin P., and Mayer, T. (2011). How do different exporters react to exchange rate changes? Theory and empirics. *Quarterly Journal of Economics* 127(1): 4437-4493.

[8] Berry, S. (1994). Estimating Discrete-Choice Models of Product Differentiation. *RAND Journal of Economics* 25(2): 242-262.

[9] Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica* 63: 841-890.

[10] Berry, S. and Pakes, A. (1993). Some applications and limitations of recent advances in empirical industrial organization: Merger analysis. *The American Economic Review* 83(2): 247-252.

[11] Bettendorf, L. and Verboven, F. (2000). Incomplete transmission of coffee bean prices: evidence from the Netherlands. *European Review of Agricultural Economics* 27(1): 1-16.

[12] Bhat, C.R. (2005). A multiple discretecontinuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological* 39(8): 679-707.

[13] Bonnet, C. and Bouamra-Mechemache, Z. (2016). Organic label, bargaining power, and profit sharing in the French fluid milk market. *American Journal of Agricultural Economics* 98(1): 113-133.

[14] Bonnet, C., Bouamra-Mechemache, Z., and Corre, T. (2018). An Environmental Tax Towards more Sustainable Food: Empirical Evidence of the Consumption of Animal Products in France. *Ecological Economics* 147(May): 48-61.

[15] Bonnet, C., Bouamra-Mechemache, Z., and Richards, T.J. (2018). Complementarity and Bargaining Power. *European Review of Agricultural Economics* 45(3): 297-331.

[16] Bonnet, C. and Dubois, P. (2010). Inference on Vertical Contracts between Manufacturers and Retailers Allowing for Non Linear Pricing and Resale Price Maintenance. *Rand Journal of Economics* 41(1): 139-164.

[17] Bonnet, C. Dubois, P., Villas-Boas, S. B., and Klapper, D. (2013). Empirical Evidence on the Role of Non Linear Wholesale Pricing and Vertical Restraints on Cost Pass-Through. *Review of Economics and Statistics* 95(2): 500-515.

[18] Bonnet, C. and Réquillart, V. 2011. Does the EU Sugar Policy Reform Increase Added Sugar Consumption? An Empirical Evidence on the Soft Drink Market. *Health Economics* 20(9): 1012-1024.

[19] Bonnet, C. and Réquillart, V. (2013a). Impact of Cost Shocks on Consumer Prices in Vertically Related Markets: The Case of the French Soft Drink Market. *American Journal of Agricultural Economics* 95: 1088-1108.

[20] Bonnet, C. and Réquillart, V. (2013b). Tax Incidence with Strategic Firms on the Soft Drink Market. *Journal of Public Economics* 106: 77-88.

[21] Bonnet, C. and Réquillart, V. (2015). Price Transmission in Food Chains: The Case of the Dairy Industry. *Food Price Dynamics and Price Adjustment in the EU*. Edited by Steve McCorriston, Oxford University Press, Chapter 4: 65-101.

[22] Bonnet, C. and Villas-Boas, S. B. (2016). An Analysis of Asymmetric Consumer Price Responses and Asymmetric Cost Pass-Through in the French Coffee Market. *European Review of Agricultural Economics* 43(5): 781-804.

[23] Brooks, K. and Lusk, J.L. (2010). Stated and revealed preferences for organic and cloned milk: combining choice experiment and scanner data. *American Journal of Agricultural Economics* 92(4): 1229-1241.

[24] Caillavet, F., Fadhuile, A. and Nichle, V. (2016). Taxing animal-based foods for sustainability: environmental, nutritional and social perspectives in France. *European Review of Agricultural Economics* 43(4): 537-560.

[25] Christensen, L.R., Jorgenson, D.W. and Lau, L.J. (1975). Transcendental logarithmic utility functions. *The American Economic Review* 65(3):367-383.

[26] Deaton, A. and Muellbauer, J. (1980a). An Almost Ideal Demand System. *The American Economic Review* 70(3): 312-326.

[27] Deaton, A. and Muellbauer, J. (1980b). *Economics and Consumer Behavior.* Cambridge University Press.

[28] Dhar, T. and Foltz, J.D. (2005). Milk by any other name consumer benefits from labeled milk. *American Journal of Agricultural Economics* 87(1): 214-228.

[29] Draganska, M., Klapper, D., and Villas-Boas, S. B. (2010). A Larger Slice or a Larger Pie? An Empirical Investigation of Bargaining Power in the Distribution Channel. *Marketing Science* 29(1): 57-74.

[30] Dubé, J.P. (2004). Multiple discreteness and product differentiation: Demand for carbonated soft drinks. *Marketing Science* 23(1): 66-81.

[31] Edjabou, L.D. and Smed, S. (2013). The effect of using consumption taxes on foods to promote climate friendly diets–The case of Denmark. *Food Policy* 39: 84-96.

[32] Goldberg, P.K. and Hellerstein, R. (2008). A structural approach to explaining incomplete exchange-rate pass-through and pricing-to-market. *American Economic Review* 98(2): 423-29.

[33] Goldberg, P. and Hellerstein, R. (2013). A structural approach to identifying the sources of local currency price stability. *Review of Economic Studies* 80(1): 175-210.

[34] Griffith, R., Nesheim, L., and O'Connell, M. (2018). Income effects and the welfare consequences of tax in differentiated product oligopoly, *Quantitative Economics* 9: 305-341.

[35] Hausman, J. A. (1997). Valuation of New Goods under Perfect and Imperfect Competition. *The Economics of New Goods*. Edited by Bresnahan, T. F and R. J. Gordon, The University of Chicago Press, Chicago/London.

[36] Hausman, J., Leonard, G. and Zona, J.D. (1994). Competitive analysis with differenciated products. *Annales d'Economie et de Statistique* 1994: 159-180.

[37] Hellerstein, R. and Villas-Boas, S. B. (2010). Outsourcing and Pass-Through. *Journal of International Economics* 81: 170-183.

[38] Kiesel, K. and Villas-Boas, S.B. (2007). Got organic milk? Consumer valuations of milk labels after the implementation of the USDA organic seal. *Journal of Agricultural and Food Industrial Organization* 5(1): Article 4.

[39] Kim, J., Allenby, G.M., and Rossi, P.E. (2002). Modeling consumer demand for variety. *Marketing Science* 21(3): 229-250.

[40] Kim, D. and Cotterill, R.W. (2008). Cost passthrough in differentiated product markets: The case of US processed cheese. *The Journal of Industrial Economics* 56(1):32-48.

[41] Kuksov, D. and Villas-Boas, J. M. (2008). Endogeneity and Individual Consumer Choice. *Journal of Marketing Research* 45(6): 702-714.

[42] Lewbel, A. and Pendakur, K. (2009). Tricks with Hicks: The EASI Demand System. *American Economic Review* 99(3): 827-63.

[43] McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*. Edited by P. Zarembka, Academic Press: New York, 1974: 105:142.

[44] McFadden, D. (1978). Modeling the choice of residential location. *Transportation Research Record* 673:72-77.

[45] McFadden, D. and Train, K.E. (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics* 15: 447-470.

[46] Meza, S. and Sudhir, K. (2010). Do private labels increase retailer bargaining power? *Quantitative Marketing and Economics* 8(3): 333-363.

[47] Misra S., Mohanty S. (2006). Estimating bargaining games in distribution channels. Working paper, University of Rochester.

[48] Nakamura E. (2008). Pass-Through in Retail and Wholesale. *American Economic Review* 98: 430437.

[49] Nakamura E., and Zerom, D. (2010). Accounting for Incomplete Pass-Through. *Review of Economic Studies* 77(3): 11921230.

[50] Nevo, A. (2000a). A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand. *Journal of Economics & Management Strategy* 9(2000):513-548.

[51] Nevo, A. (2000b). Mergers with Differentiated Products: The case of the Ready-to-Eat Cereal Industry. *RAND Journal of Economics*, 31 (1): 395-421.

[52] Petrin, A. and Train, K. (2010). A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research* 47(1): 313.

[53] Revelt, D., and Train, K. (2000). Customer-specific Taste Parameters and Mixed Logit: Households Choice of Electricity Supplier. Working Paper No. E00-274, Department of Economics, University of California, Berkeley.

[54] Richards, T.J., Acharya, R.N. and Molina, I. (2011). Retail and wholesale market power in organic apples. *Agribusiness* 27(1): 62-81.

[55] Richards, T.J., Gómez, M.I. and Pofahl, G. (2012). A multiple-discrete/continuous model of price promotion. *Journal of Retailing* 88(2):206-225.

[56] Rivers, D. and Vuong, Q. (2002). Model selection tests for nonlinear dynamic models. *The Econometrics Journal* 5(1): 1-39.

[57] Rosse, J.N. (1970). Estimating cost function parameters without using cost data: Illustrated methodology. *Econometrica: Journal of the Econometric Society* 1970: 256-275.

[58] Slade, M.E. (2004). Market power and joint dominance in UK brewing. *The Journal of Industrial Economics* 52(1): 133-163.

[59] Smith, R.J. (1992). Non-nested tests for competing models estimated by generalized method of moments. *Econometrica: Journal of the Econometric Society* 1992: 973-980.

[60] Song, I. and Chintagunta, P.K. (2006). Measuring cross-category price effects with aggregate store data. *Management Science* 52(10): 1594-1609.

[61] Stone, R. (1954). Linear expenditure systems and demand analysis: an application to the pattern of British demand. *The Economic Journal* 64(255): 511-527.

[62] Train, K.E. (2003). Discrete Choice Methods with Simulation. Cambridge: University Press, 2003.

[63] Villas-Boas, S.B. (2007a). Vertical Relationships Between Manufacturers and Retailers: Inference With Limited Data. *The Review of Economic Studies* 74(2): 625-52.

[64] Villas-Boas, S.B. (2007b). Using Retail Scanner Data for Upstream Merger Analysis. *Journal of Competition Law and Economics* 3(4): 689-715.

[65] Villas-Boas, S. B. and Hellerstein, R. (2006). Identification of Supply Models of Retailer and Manufacturer Oligopoly Pricing. *Economics Letters* 90(1): 132-40.

[66] Villas-Boas, J. M. and Zhao, Y. (2005). Retailer, Manufacturers, and Individual Consumers: Modeling the Supply Side in the Ketchup Marketplace. *Journal of Marketing Research* 42: 83-95.

[67] Vuong, Q.H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society* 1989: 307-333.

[68] Werden, G. J. and Froeb, L. M. (1994). The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy. *The Journal of Law, Economics, & Organization* 10(2): 407-26.